NMR Database of Lignin and Cell Wall Model Compounds

Structure Index

Contents

Electronic Version: If you are viewing this on your computer with Adobe's Acrobat Reader (or in NetScape etc.) the Hyperlinks are all active. Click on a Content item, or a structure in the Structure Index, to go to that item.

NMR Database

of

Lignin and Cell Wall Model Compounds

Sally Ralph and **Larry Landucci**
US Forest Products Laboratory
USDA Forest Service
One Gifford Pinchot Drive
Madison, WI 53705
Email: saralph@facstaff.wisc.edu
Email: lllanduc@facstaff.wisc.edu

John Ralph
US Dairy Forage Research Center
USDA-Agricultural Research Service
1925 Linden Drive West
Madison, WI 53706
Email: jralph@facstaff.wisc.edu

Introduction

This database was created and is administered as a cooperative effort between the US Forest Products Laboratory and the US Dairy Forage Research Centers in Madison, Wisconsin. It was designed to provide a coherent, single source of NMR data of lignin model compounds as well as compounds modeling similar structures in grasses and other forage plants.

Database Formats/Versions

The database exists in four different formats: an interactive Hypercard© stack for the Macintosh® computer, a FileMaker Pro© database for cross-platform use, an Adobe© cross-platform portable document format (pdf) file for viewing and printing using Adobe's free Adobe Acrobat Reader, and a hardcopy version derived from the FileMaker Pro database. The first three versions are available for downloading over the internet from the Dairy Forage Research Center web site:

http://www.dfrc.wisc.edu

The hardcopy is expensive to produce but is available by request from the authors at the Forest Products Laboratory to researchers who do not have computer access.

Data Collection

In general ^{13}C NMR data were collected in three common deuterated solvents (acetone, chloroform and dimethyl sulfoxide) for each compound. The ^{1}H NMR data was obtained in one solvent only. A standard set of acquisition parameters was used to

acquire and process the spectra to keep the data as uniform and constant as possible. Those compounds with an index number less than 1000 were run on a Bruker 250 MHz spectrometer at FPL and those compounds with an index number between 1000 and 10,000 were run at the DFRC on a Bruker 360 MHz instrument. The order of the compounds in the database reflects their arrival at the spectrometer rather than a preordained plan. Search routines for the software versions allow grouping the compounds with similar traits, whereas the structure index is most useful for the hardcopy and pdf versions. The inclusion of many analogous series of structures with small structural differences allows calculation of substituent effects that are invaluable for chemical shift predictions of structures not included in the database.

Assignments

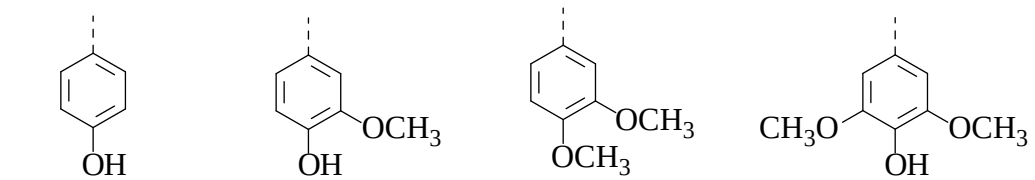
The chemical shift assignments for most of the compounds were made by comparison with other compounds, literature values and in some cases other NMR experiments such as long and short range C-H correlations, COSY and DEPT. Every effort was made to correctly assign the chemical shifts; however, limited time and resources precluded confirming the shifts for many of the compounds. The shifts are reported to the second decimal place only to distinguish very close shifts; comparisons between spectra are practical only within ± 0.1 ppm. The authors would greatly appreciate any corrections on misassignments.

Compound Sources

The compounds themselves came from many sources; in house collections, syntheses and donations from other researchers for which we are grateful. The source of the compounds is often given in the “Notes” field along with other pertinent data. The intensities of the individual chemical shift signals are used for the line plots generated by the Hypercard program but they are also useful in the hardcopy version for comparison with spectra.

Naming Conventions

This database was originally intended as an aid for the assignment of chemical shifts for wood and plant lignin NMR spectra. The trivial names used throughout are well known to wood chemists as is the numbering system. We have attempted to include more formal chemical names for many of the compounds and these were obtained using Beilstein’s Autonom© program. The chemical names for the larger 3 and 4 ring models became so cumbersome that the authors employed an abbreviated system to identify both the moieties involved as well as the linkages between the moieties. Examples of the naming, numbering and linking conventions used are given below.



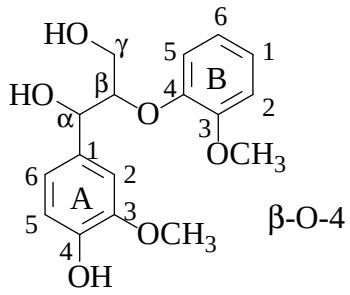
4-hydroxyphenyl (*p*-coumaryl) guaiacyl veratryl syringyl

Fig. 1. Trivial names for the common aromatic ring substitutions.

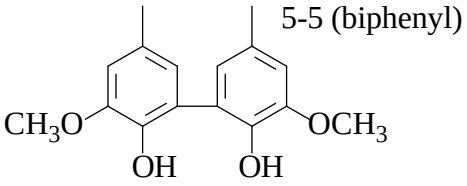
The naming of the larger oligomer lignin models uses a combination of upper case letters to describe the ring structure and lower case letters and numbers to describe the type of linkage between the rings.

Table 1
Terminology for Structural Entities

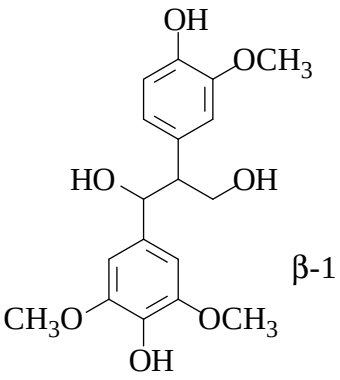
Entity	Abbreviation
guaiacyl ring	G
syringyl ring	S
p-coumaryl ring	H
α–O–4 linkage	a
β–O–4 linkage	b
β–5 (phenylcoumaran)	c
β–1 linkage	b1
β–β (resinol)	r
5-5 (biphenyl)	5,5
coniferyl alcohol end unit	CA
sinapyl alcohol end unit	SA
p-coumaryl alcohol end unit	HA
ferulic acid end unit	FA
erythro	e
threo	t



G-b-G

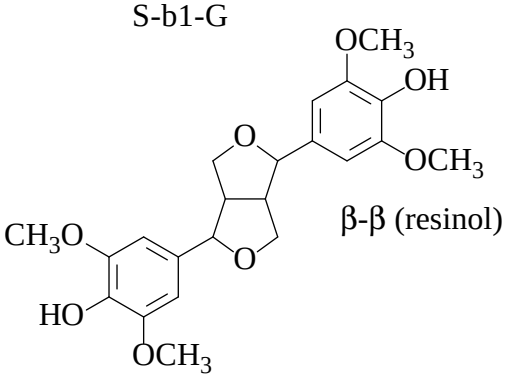


G-5,5-G

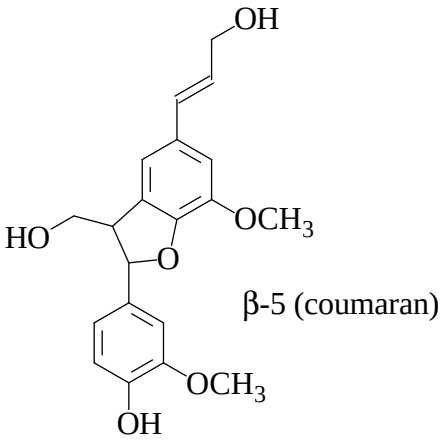


S-b1-G

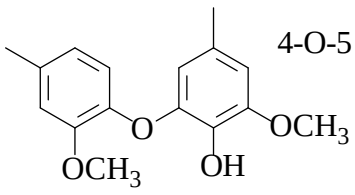
Fig. 2. Examples of linkages and abbreviated names.



S-r-S

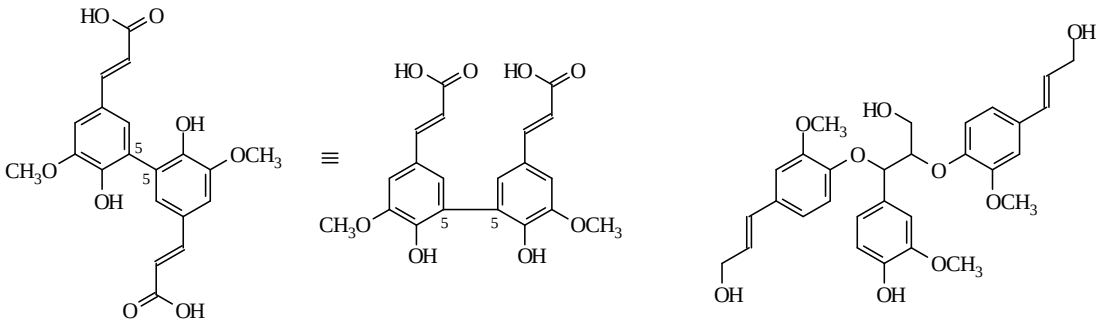


G-c-CA



G-4-O-5-G

With this convention the name FA-5,5-FA would represent a diferulic acid biphenyl structure. The trimer CA-a-G-b-CA would be a guaiacyl unit with two coniferyl alcohol end groups etherified at the alpha and beta positions.



FA-5,5-FA **CA-a-G-b-CA**

Structure Index

The structure index is arranged based upon the number of rings in the structure. Where possible the structures are also arranged by ring type such as guaiacyl, syringyl etc. The number under the structure refers to the index number at the top of the data sheet. An asterisk after a number indicates the acetylated analog of that compound. In some cases only the acetylated compound is included.

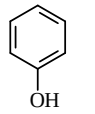
Future Plans

We hope to continue adding to and improving this database. Regular updates will be made to the database to keep the on-line sources current. We will also make every effort to keep those researchers with hardcopy versions supplied with more pages. This database was written and prepared by U.S. Government employees on official time, and it is therefore in the public domain and not subject to copyright. Please feel free to contact the authors with suggestions or questions.

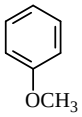
Acknowledgments

The authors gratefully acknowledge the many generous contributions made by others at both labs towards this database; William Landucci for software development, Martin Wesolowski and Kolby Hirth, NMR spectroscopists, and the frequent donation of compounds from Mike Mozuch, Noritsugu Terashima, Stéphane Quideau, Rich Helm, Jamie Milhaupt and Susana Luque. The authors also acknowledge partial support by the USDA-National Research Initiative Competitive Grants Program, Improved Utilization of Wood and Fiber Section, award #94-03465.

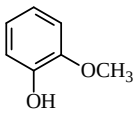
Monomers



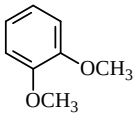
38,190*



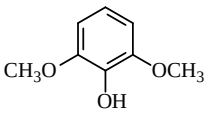
54



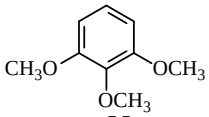
39,176*



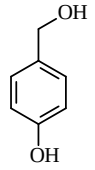
56



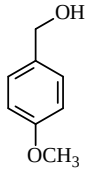
40,189*



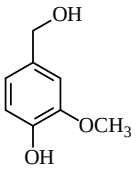
55



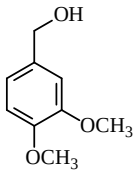
4,24*



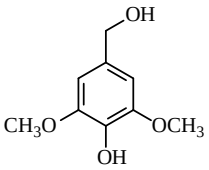
37,205*



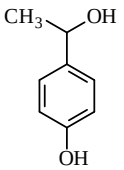
19,200*



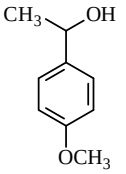
5,143*



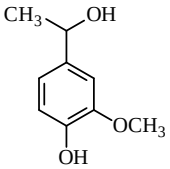
16,201*



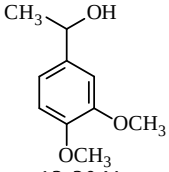
10,23*



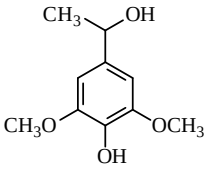
209,210*



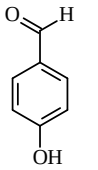
11,202*



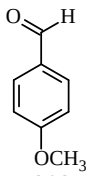
13,204*



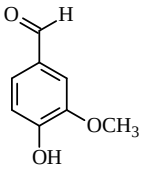
12,195*



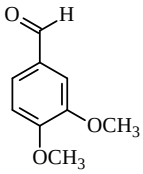
14,207*



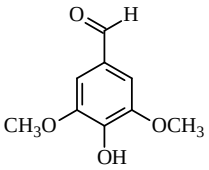
203



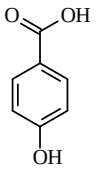
15,174*



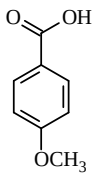
57



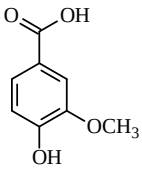
41,206*



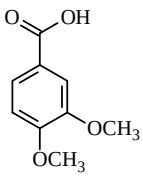
58,208*



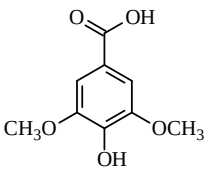
20



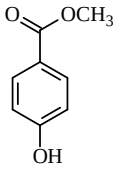
42,192*



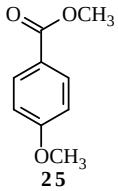
44



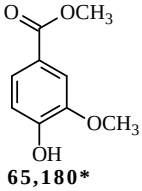
43,175*



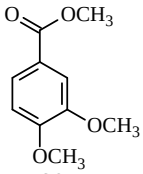
18,17*



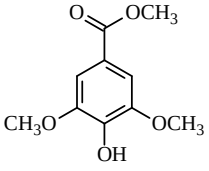
25



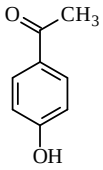
65,180*



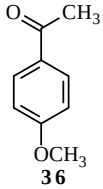
69



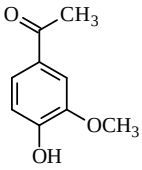
64,191*



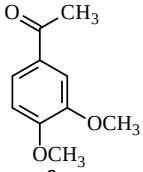
6,173*



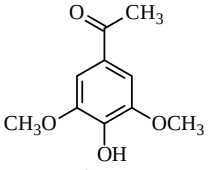
36



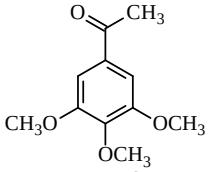
7,103*



9

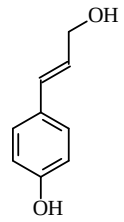


8,177*

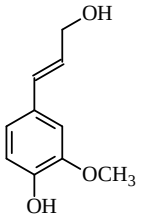


156

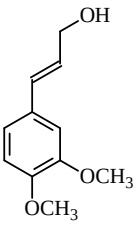
Monomers con't



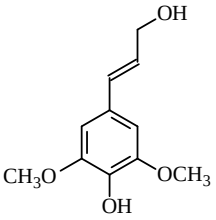
151,2003



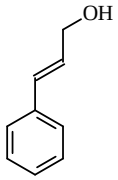
87,2001,142*



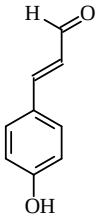
145,159*



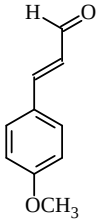
2002,211*



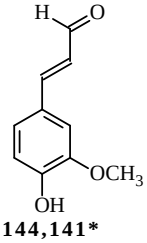
95



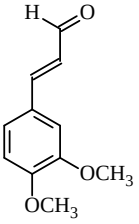
152



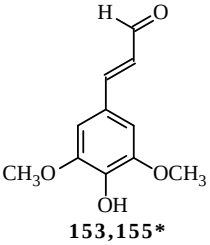
215



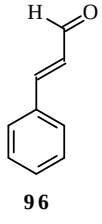
144,141*



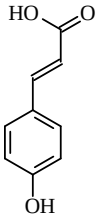
150



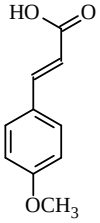
153,155*



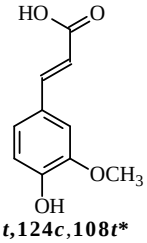
96



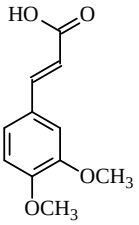
45,129*



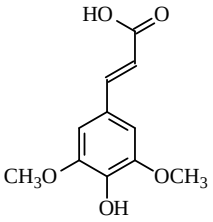
49



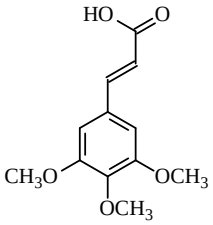
48*t*,124*c*,108*t*



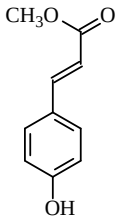
47



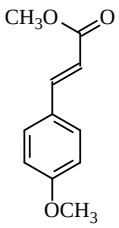
46,178*



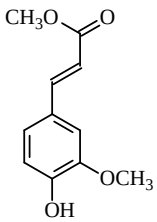
63



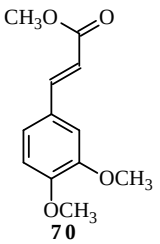
61,148*



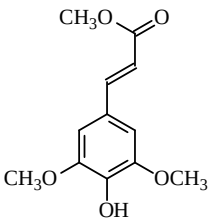
59



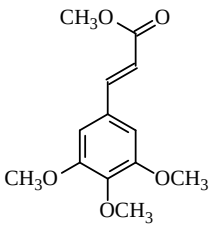
68,149*



70

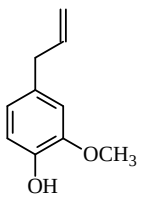


60,196*

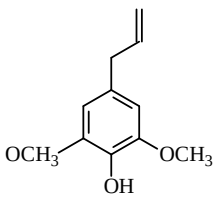


62

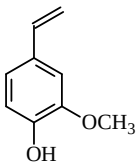
Misc. Monomers



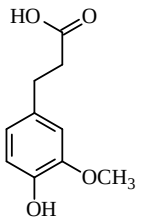
113,115*



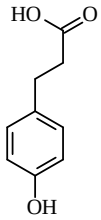
114,116*



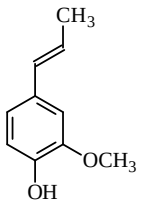
136



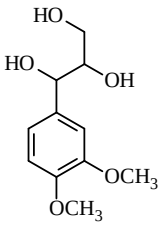
35



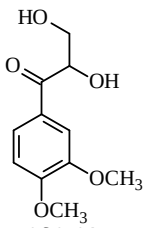
127,128*



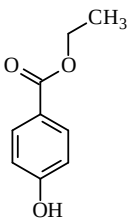
112,111*



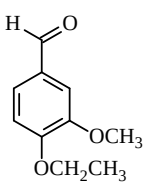
160



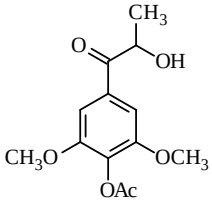
161,1*



94

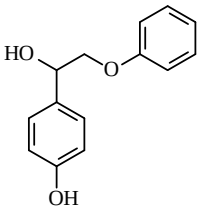


167

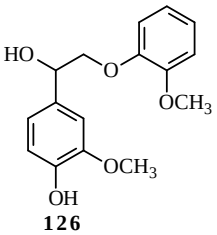


130

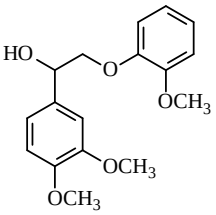
β-O-4 Dimers, 2 Carbon Sidechain



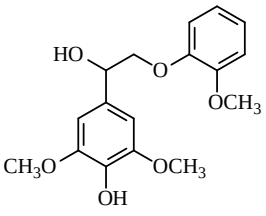
125



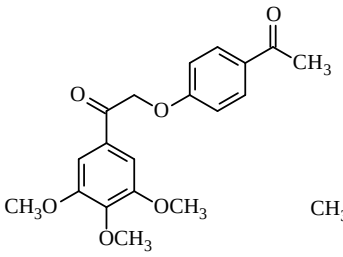
126



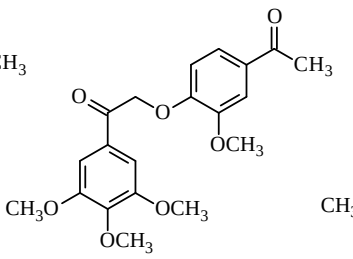
147



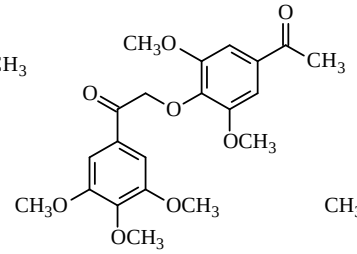
50



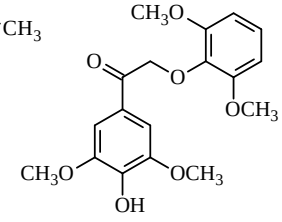
51



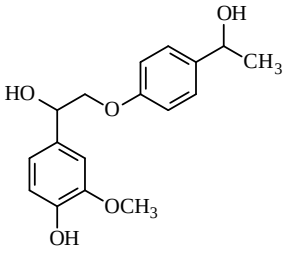
52



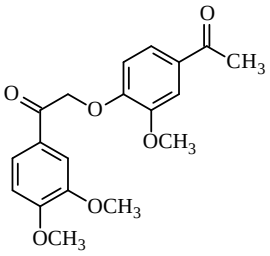
53



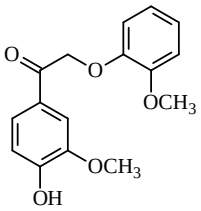
22,26*



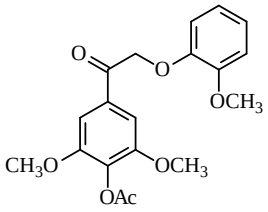
132,138*



146

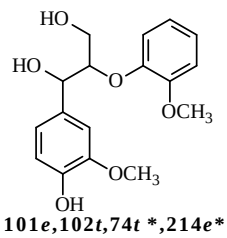


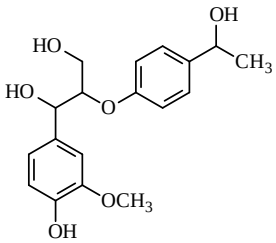
154,1027*

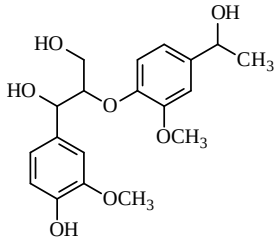


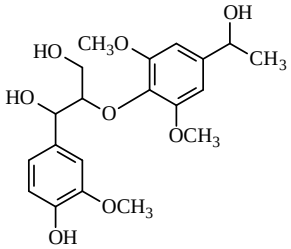
27

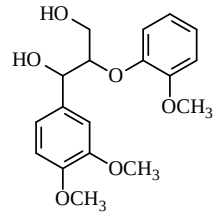
β-O-4 Dimers, 3 Carbon Sidechain

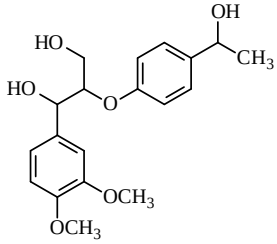

101*e*,102*t*,74*t* *,214*e**

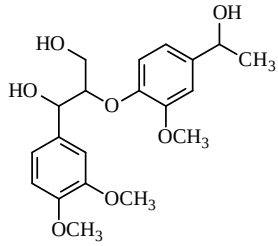

133,137*

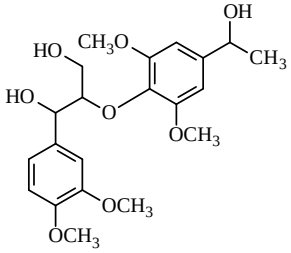

131,140*

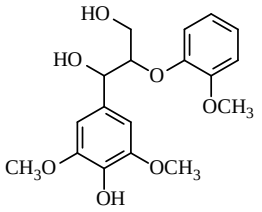

134,139*

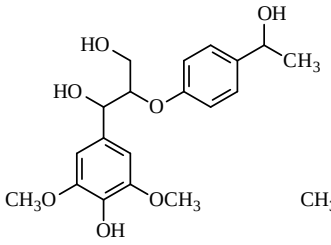

104*t*,105*e*,3*e* *

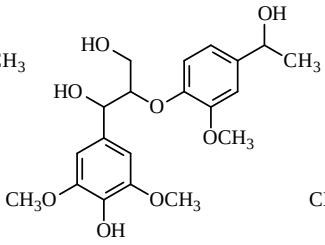

31,32*

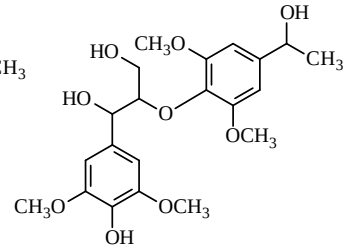

34,33*

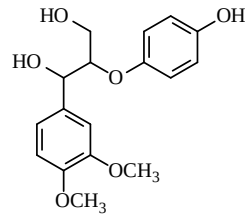

2,29*

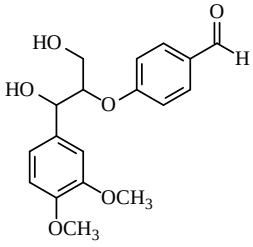

90,99*

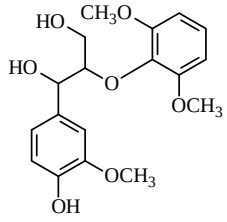

91,100*

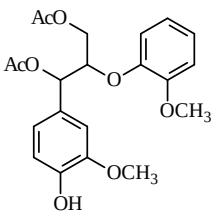

88,97*

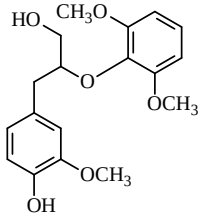

89,98*

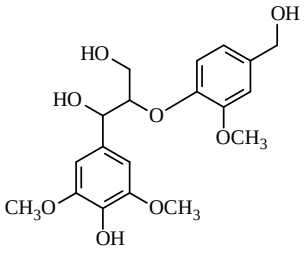

166

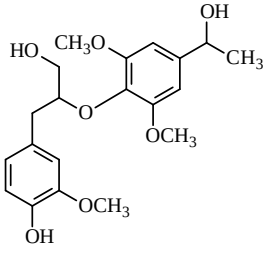

168


179

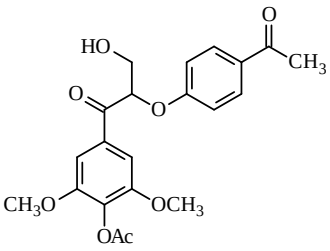

1029*t*,1030*e*

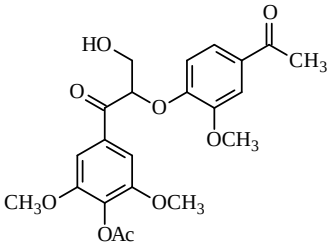

217,218*

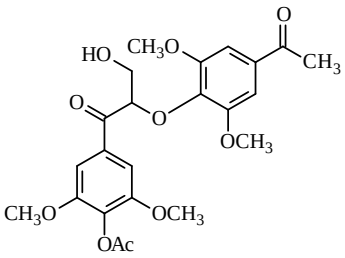

212,213*

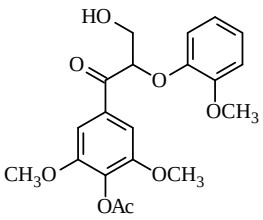

135

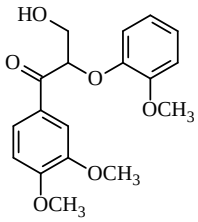
More β-O-4 Dimers, 3 Carbon Sidechain

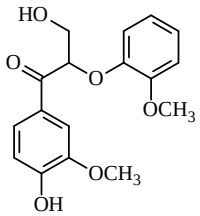

80

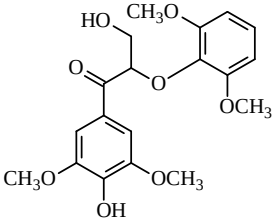

81

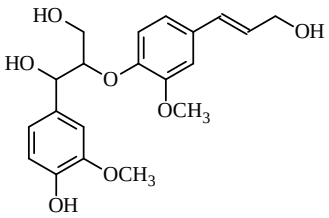

79

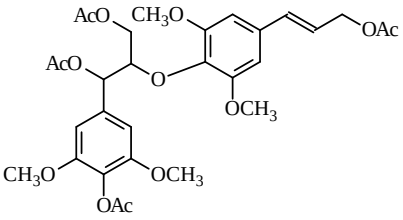

92,2067

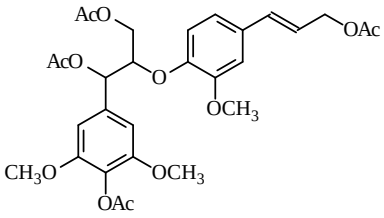

110,163*

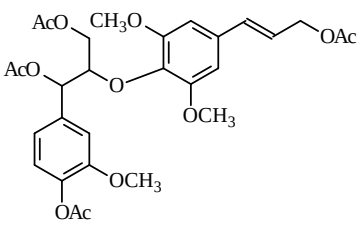

106,162*,1028* (monoacetate)

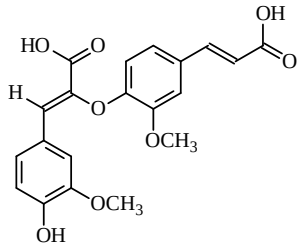

21,2082* (monoacetate)

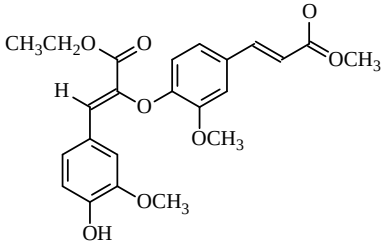

2013*t*,2014*e*,2012*e* *,2011*t**

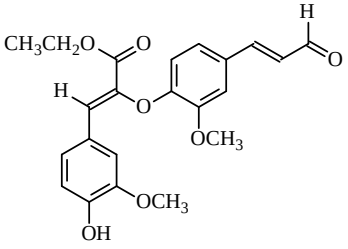

185

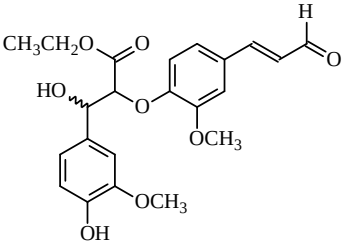

186

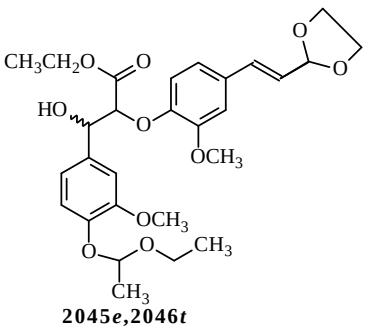

188


2040

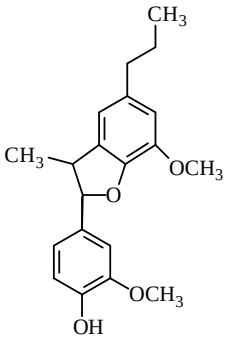

2041


2042

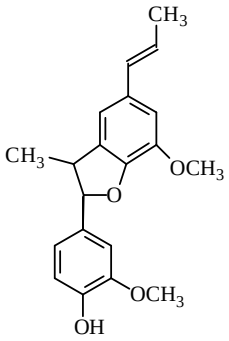

2043*e*,2044*t*


2045*e*,2046*t*

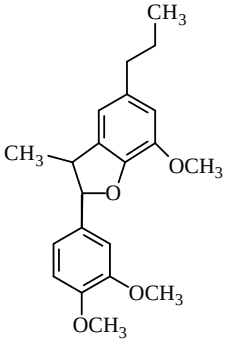
β-5 Dimers



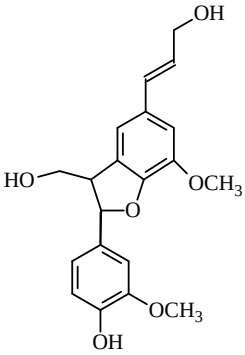
66



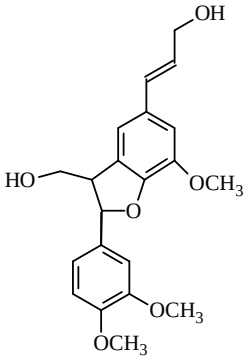
67



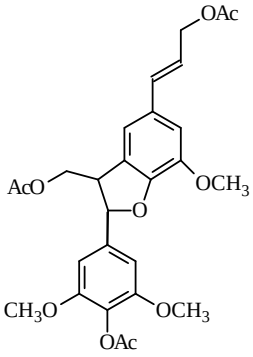
73



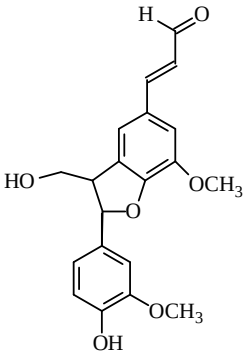
2004,2005*



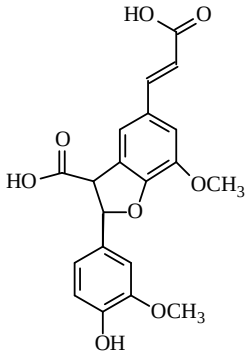
2051



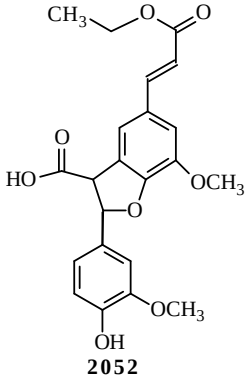
187



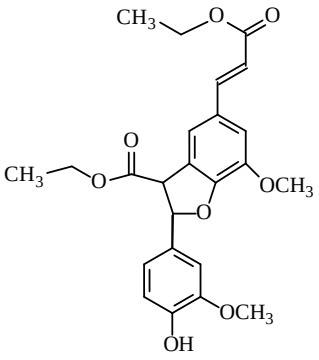
2021



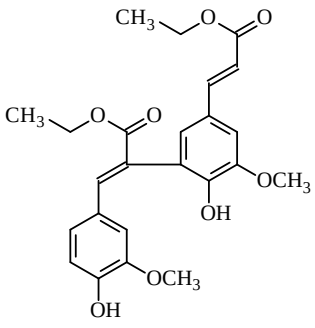
2018



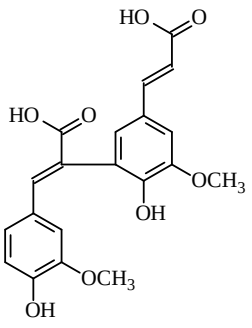
2052



2006

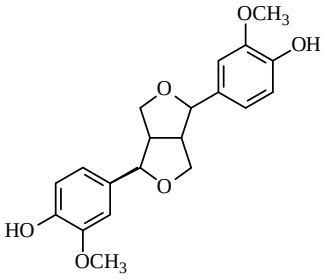


2019

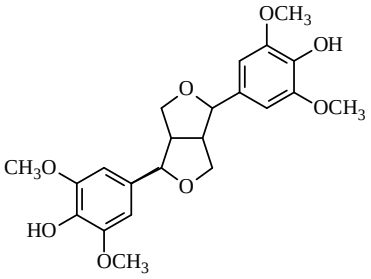


2054

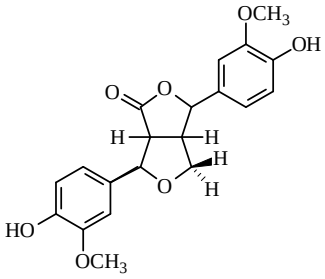
β-β **Dimers**



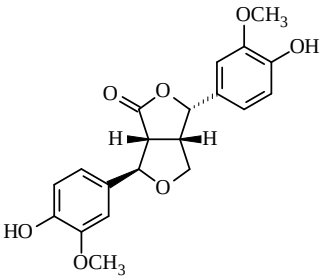
2020,109*



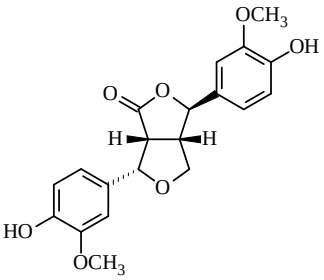
117,123*



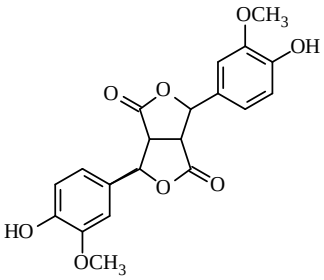
2029,2030*



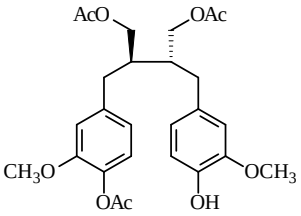
2028



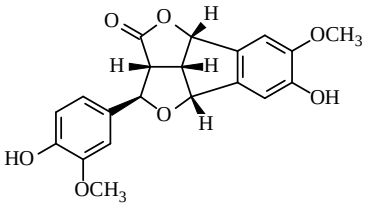
2027



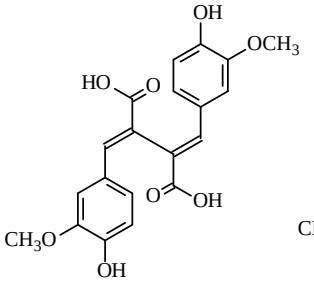
2008



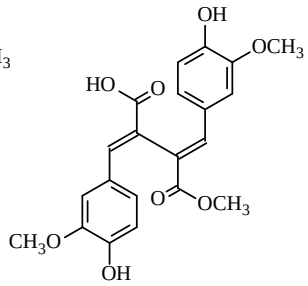
2070



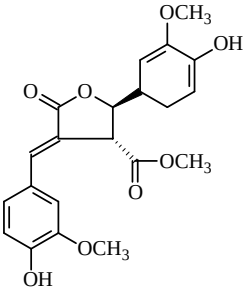
2026



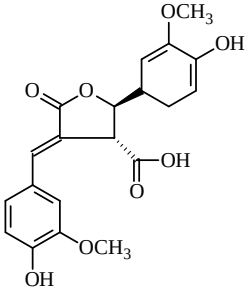
2031



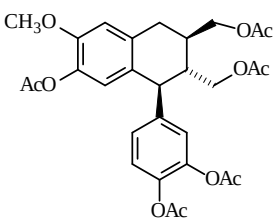
2032



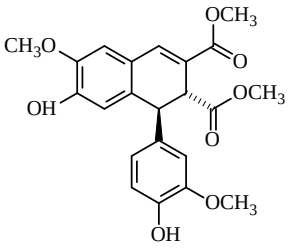
2034



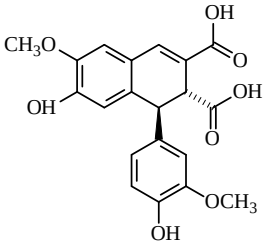
2033



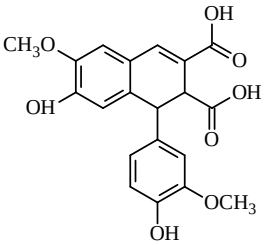
2069



2035

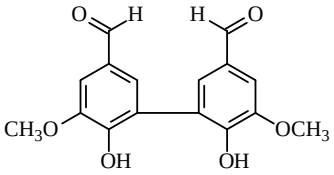


2062

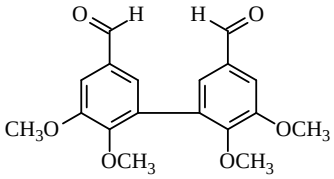


2036

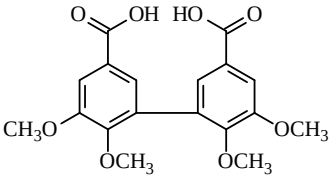
5-5 Dimers



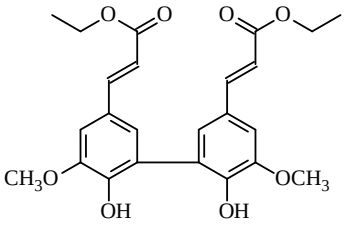
28



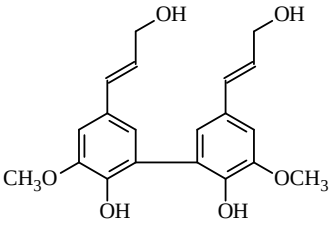
71



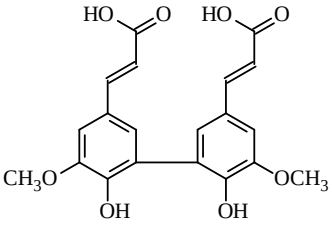
72



2057

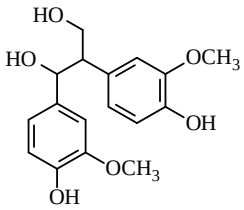


2058



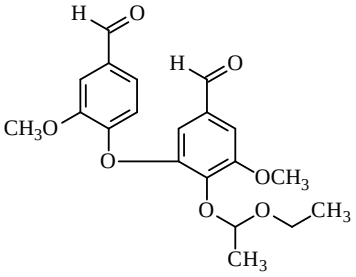
2056

β-1 Dimer

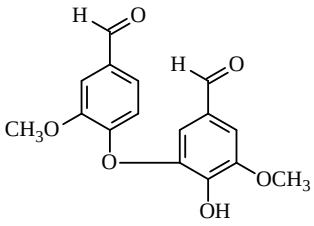


171,172*

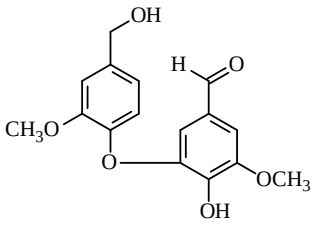
4-O-5 Dimers and Trimer



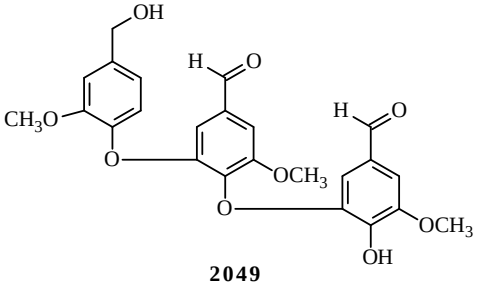
2037



2038

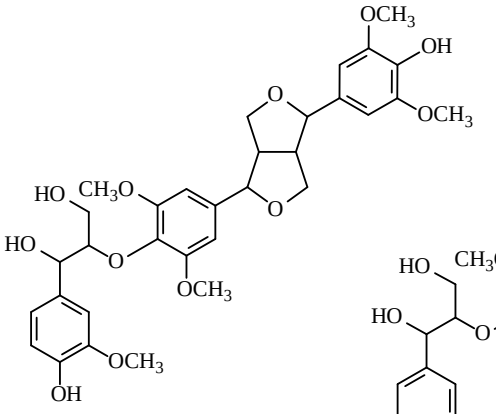


2039

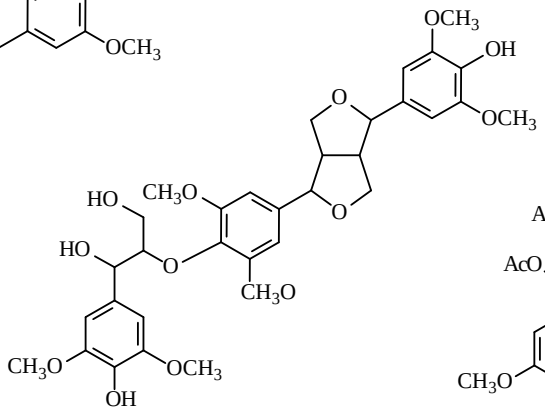


2049

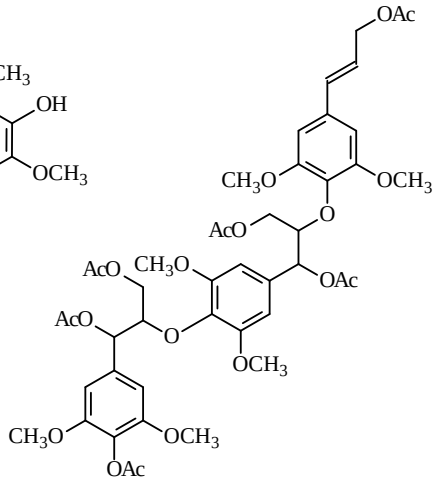
Trimers



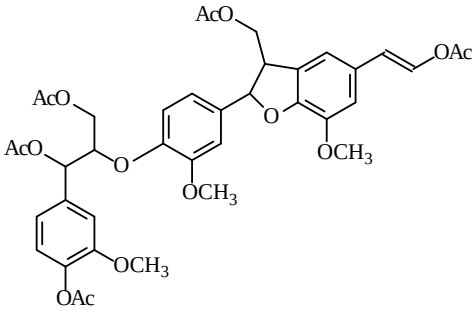
183,184*



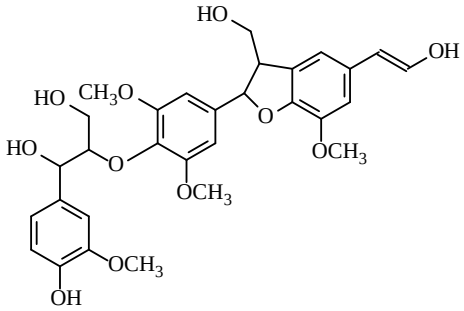
198,199*



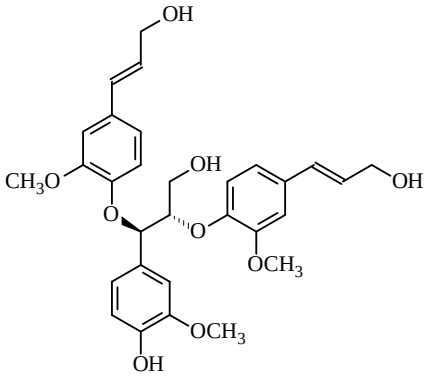
193



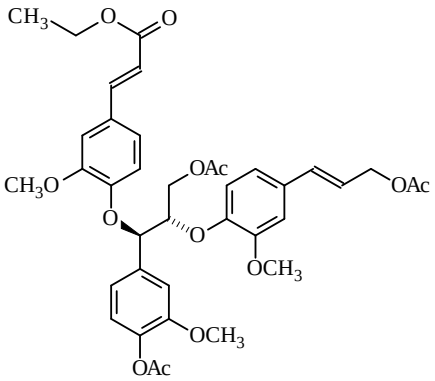
216



181,182*

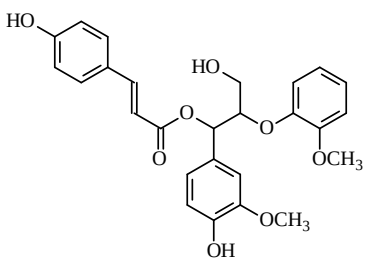


2015e,2022e*

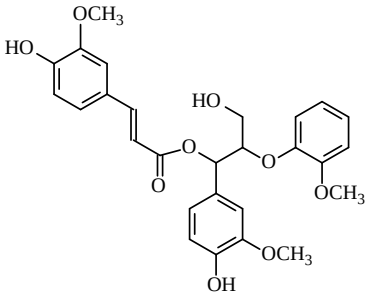


2016e

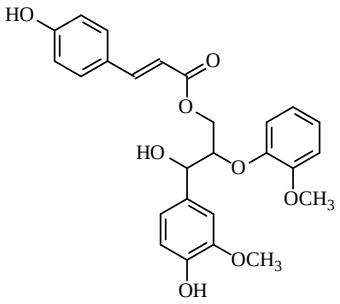
Trimers Containing Ferulic or Coumaric Acid



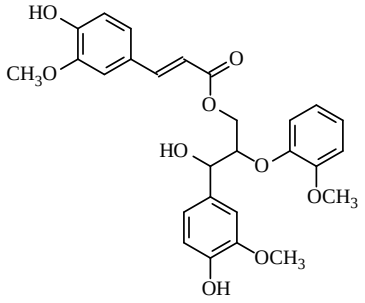
1001*t*,1002*e*,1005*t* *,1006*e* *,78*t* *,77*e* *



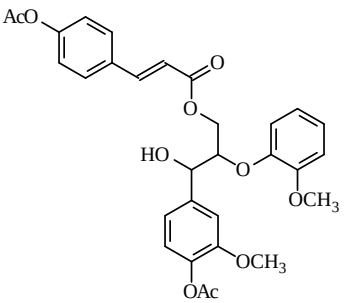
1003*t*,1004*e*,1007*t* *,1008*e* *



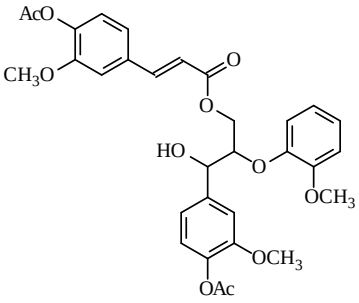
1015*t*,1016*e*,1019*t* *,1020*e* *,82*t* *,83*e* *



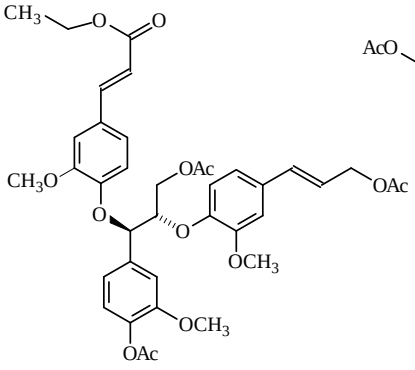
1017*t*,1018*e*,1021*t* *,1022*e* *,76*e*,75*t* *



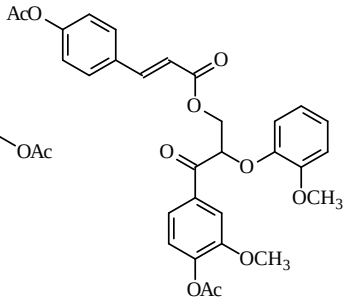
1011*t*,1012*e*



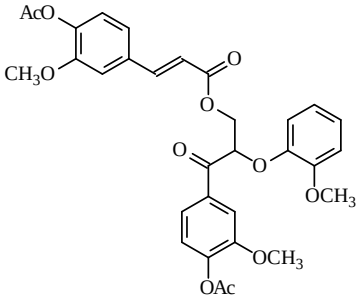
1013*t*,1014*e*



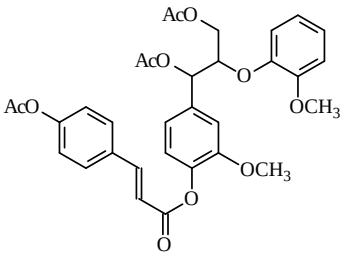
2016



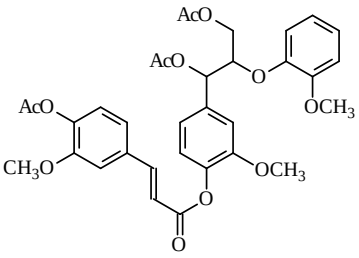
1009



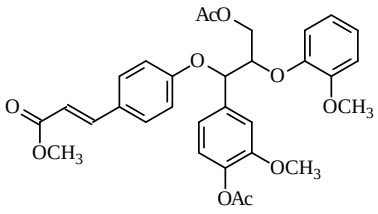
1010



1023*t*,1024*e*

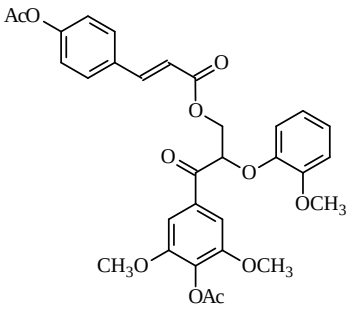


1025*t*,1026*e*,84*e*

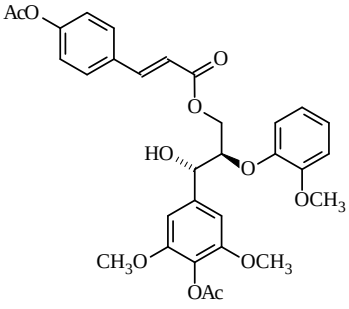


85*e*,86*t*

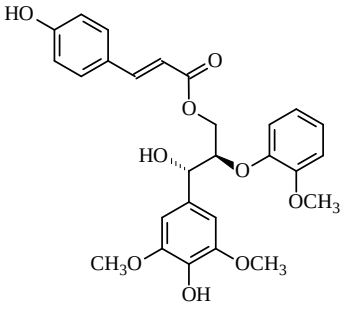
More Trimers Containing Ferulic or Coumaric Acid



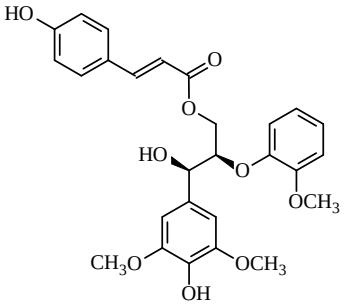
2066



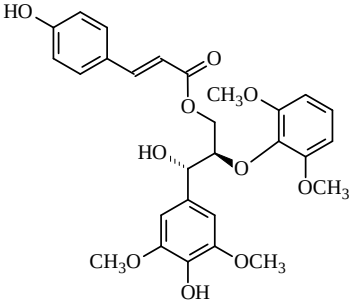
2068



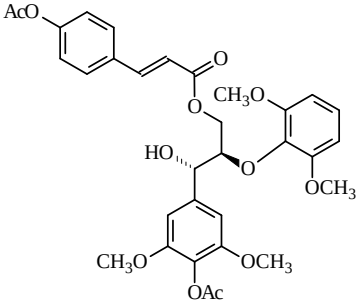
2078,2077*



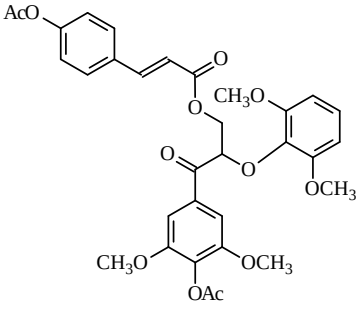
2076,2075*



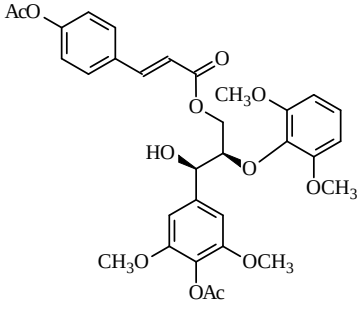
2072,2074*



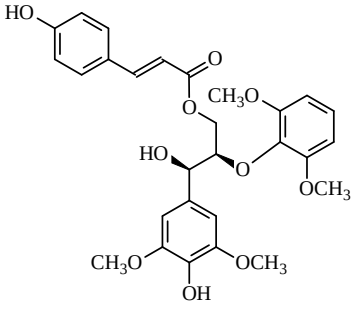
2080



2081

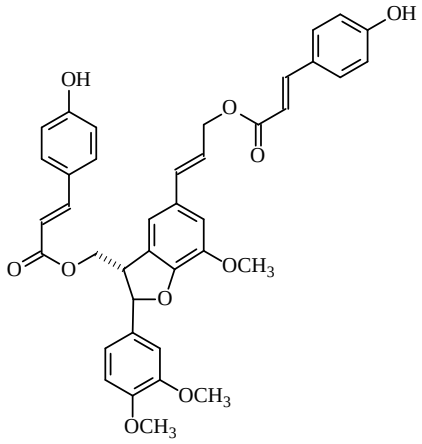


2079

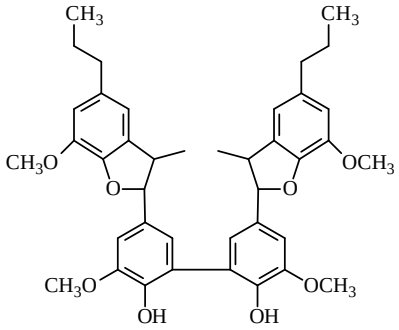


2071,2073*

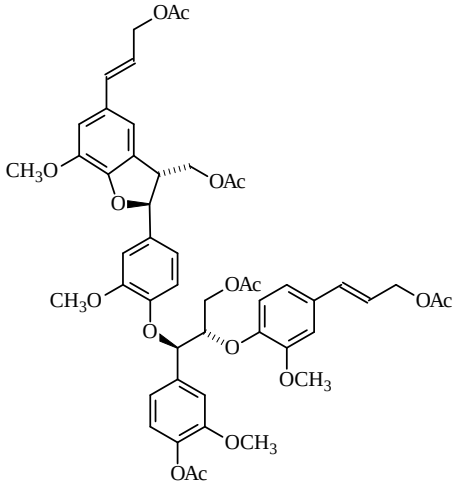
Tetramers



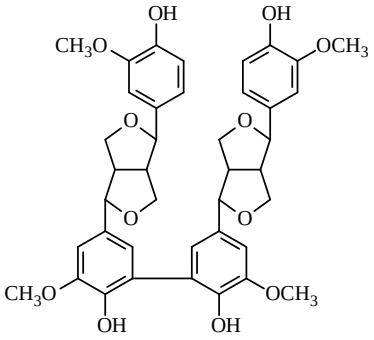
2060,2061*



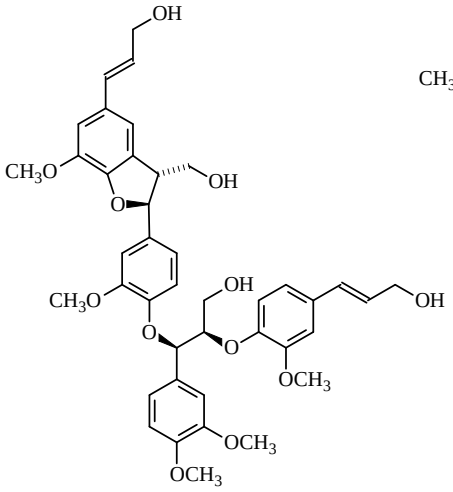
170,169*



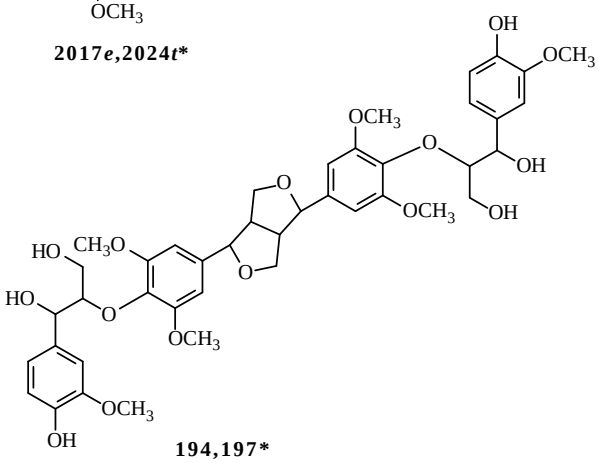
2023*e*



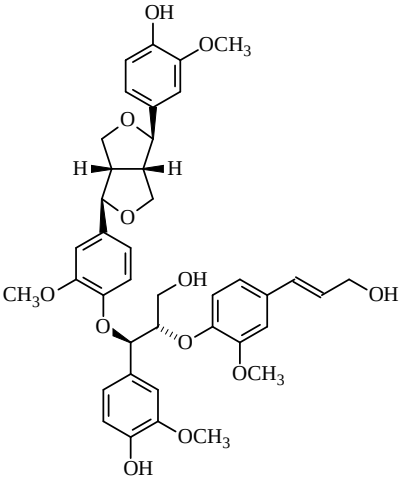
164,165*



2017*e*,2024*t**

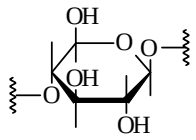


194,197*

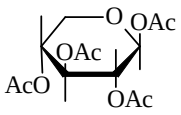


2025*e* ,2059*

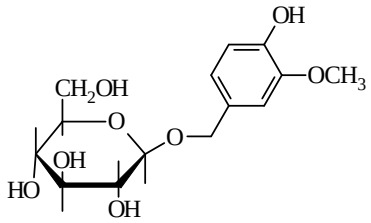
Misc. Compounds



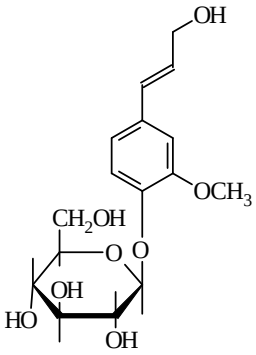
30



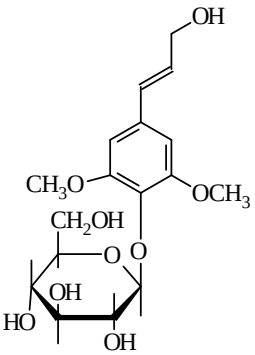
122, *a* &*b*



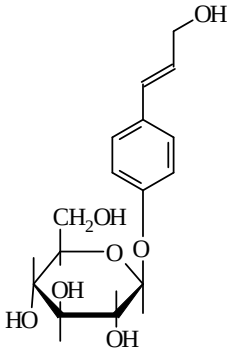
107



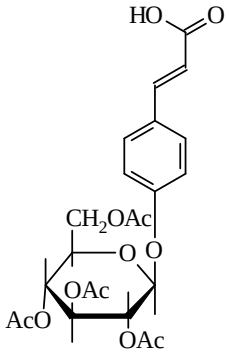
118,120*



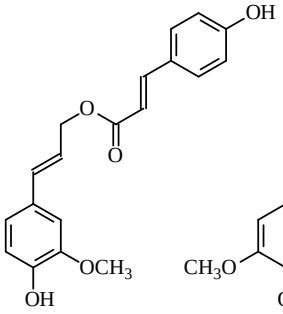
119,121*



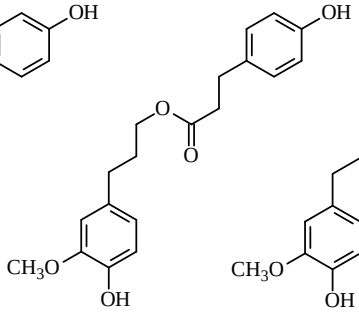
157



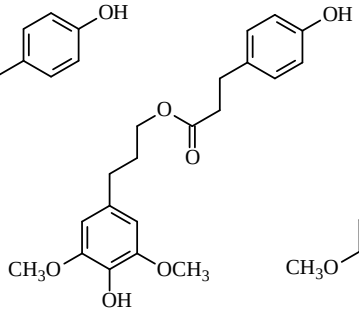
158*



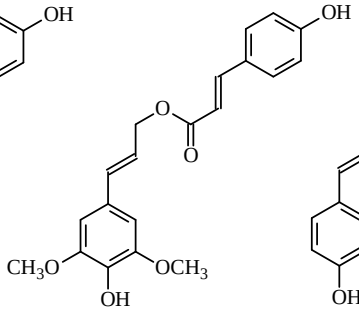
2009



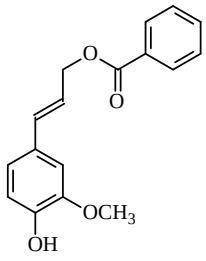
2063



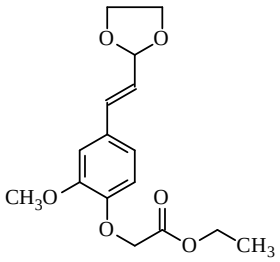
2064



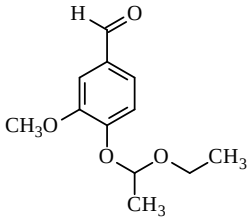
2065



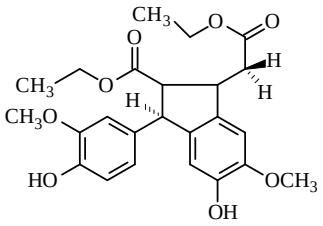
2010



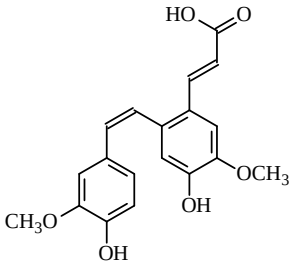
2047



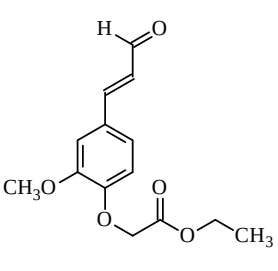
2048



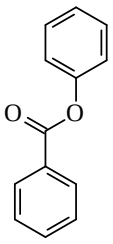
2050



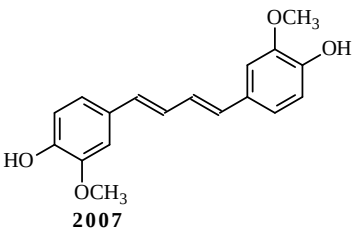
2053



2055



93



2007